3D face generation tool Candide for better face matching in surveillance video

Prepared by:
Dmitry Gorodnichy
Science & Engineering Directorate, Canadian Border Services Agency
79 Bentley Avenue, Ottawa, Ontario K1A 0L8 Canada

Eric Granger
École de technologie supérieure, Université du Québec

Stan Matwin
Erico Neves
School of Electrical Engineering and Computer Science of the University of Ottawa

Scientific Authority:
Pierre Meunier
DRDC Centre for Security Science
613-944-4367

Contract Report
DRDC-RDDC-2014-C168
July 2014

IMPORTANT INFORMATIVE STATEMENTS

PROVE-IT (FRiV) Pilot and Research on Operational Video-based Evaluation of Infrastructure and Technology: Face Recognition in Video, PSTP 03-401BIOM was supported by the Canadian Safety and Security Program (CSSP) which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada. Led by Canada Border Services Agency partners included : Royal Canadian Mounted Police, Defence Research Development Canada, Canadian Air Transport Security Authority, Transport Canada, Privy Council Office; US Federal Bureau of Investigation, National Institute of Standards and Technology, UK Home Office; University of Ottawa, Université Québec (ÉTS)

The CSSP is a federally-funded program to strengthen Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence.

Canada Border Services Agency / Agence des services frontaliers du Canada

Science and Engineering Directorate

Border Technology Division

Division Report 2014-10 (TR)
June 2014

3D face generation tool Candide for better face matching in surveillance video

Dmitry Gorodnichy,
Eric Granger,
Stan Matwin,
Erico Neves

PROTECTION • SERVICE • INTEGRITY

Canada

This page left intentionally blank

Abstract

One of the key challenges in surveillance video face screening against a Watch List is the fact that faces in surveillance video are often observed at an angle different from the angle at which the faces are captured in the Watch List. Particularly, facial images in surveillance video are normally observed at various pose angles, and from above eye level. In contrast, mugshot (reference facial images) stored in databases are regularly captured at a frontal post and at eye level, thus causing poor matching between the images. One way to overcome this problem is seen in advanced pre-processing of stored images. It is possible to synthetically generate variations of a reference facial images of target individuals at under the same conditions (e.g. pose angle) under which they will be most likely observed in a video. While several commercial tools exist, an open source library is available to generate a 3D face model from arbitrary 2D facial images. This library, called Candide, may allow academia and industry to significantly improve the matching performance of their algorithms in video surveillance applications. This report overviews this library and analyzes its suitability for the problem.

Keywords: video-surveillance, face recognition in video, instant face recognition, watch-list screening, biometrics, reliability, performance evaluation

Community of Practice: Biometrics and Identity Management

Canada Safety and Security (CSSP) investment priorities:

1. <u>Capability area:</u> P1.6 – Border and critical infrastructure perimeter screening technologies/ protocols for rapidly detecting and identifying threats.
2. <u>Specific Objectives:</u> O1 – Enhance efficient and comprehensive screening of people and cargo (identify threats as early as possible) so as to improve the free flow of legitimate goods and travellers across borders, and to align/coordinate security systems for goods, cargo and baggage;
3. <u>Cross-Cutting Objectives</u> CO1 – Engage in rapid assessment, transition and deployment of innovative technologies for public safety and security practitioners to achieve specific objectives;
4. <u>Threats/Hazards</u> F – Major trans-border criminal activity – e.g. smuggling people/ material

Acknowledgements

This work is done within the project PSTP-03-401BIOM “PROVE-IT(FRiV)” funded by the Defence Research and Development Canada (DRDC) Centre for Security Science (CSS) Public Security Technical Program (PSTP) by the following contributors:

1. **D. Gorodnichy**, Science & Engineering Directorate, Canadian Border Services Agency (CBSA),

2. **E.Granger**, École de technologie supérieure, Université du Québec.

3. **S. Matwin, E. Neves.** School of Electrical Engineering and Computer Science of the University of Ottawa.

Disclaimer

In no way do the results presented in this paper imply recommendation or endorsement by the Canada Border Services Agency, nor do they imply that the products and equipment identified are necessarily the best available for the purpose. The information presented in this report contains only the information available in public domain.

Release Notes

Context: This document is part of the set of reports produced for the PROVE-IT(FRiV) project. All PROVE-IT(FRiV) project reports are listed below.

- Dmitry Gorodnichy, Eric Granger "PROVE-IT(FRiV): framework and results". Also published in Proceedings of NIST International Biometrics Performance Conference (IBPC 2014), Gaithersburg, MD, April 1-4, 2014. Online at http://www.nist.gov/itl/iad/ig/ibpc2014.cfm.
- Dmitry Gorodnichy, Eric Granger, "Evaluation of Face Recognition for Video Surveillance". Also published in Proceedings of NIST International Biometric Performance Conference (IBPC 2012), Gaithersburg, March 5-9, 2012. Online at http://www.nist.gov/itl/iad/ig/ibpc2012.cfm.
- E. Granger, P.Radtke, and D. Gorodnichy, "Survey of academic research and prototypes for face recognition in video,"
- D. Gorodnichy, E.Granger, and P.Radtke, "Survey of commercial technologies for face recognition in video,"
- E. Granger and D. Gorodnichy, "Evaluation methodology for face recognition technology in video surveillance applications,'
- D. Gorodnichy, E. Granger, E. Choy, W. Khreich, P.Radtke, J. Bergeron, and D. Bissessar, "Results from evaluation of three commercial off-the-shelf face recognition systems on Chokepoint dataset,"
- S. Matwin, D. Gorodnichy, and E. Granger, "Using smooth the ROC method for evaluation and decision making in biometric systems,"
- D. Gorodnichy, E. Granger, S. Matwin, E. Neves "3D face generation tool Candide for better face matching in surveillance video,"
- E. Neves, S. Matwin, D. Gorodnichy, and E. Granger, "Evaluation of different features for face recognition in video,"

The PROVE-IT(FRiV) project took place from August 2011 till March 2013. This document was drafted and discussed with project partners in March 2013 at the Video Technology for National Security (VT4NS) forum. The final version of it was produced in June 2014.

Typesetting: All tabulated content in this report was produced automatically using LATEX content for improved source control, flexibility and maintainability. The report contains automatically generated hyper-link references and table of contents for easier navigation and reading on-line.

Contact: Correspondence regarding this report should be directed to DMITRY dot GORODNICHY at CBSA dot GC dot CA.

dmitry.gorodnichy@cbsa.gc.ca

Contents

List of Figures

1 Introduction

As discussed in [1], one of the key challenges in surveillance video face screening against a Watch List is the fact that faces in surveillance video are often observed at an angle different from the angle at which the faces are captured in the Watch List. Particularly, facial images in surveillance video are normally observed from above eye level, whereas mugshot facial images stored in databases are regularly captured at eye level, thus causing poor matching between the images. A way to overcome the problem is seen in advanced preprocessing of stored images to synthetically generate the facial images of target individuals in the poses under which they will be most likely observed in video. While there are several commercial tools to do that such as the ones developed by CMU [1], Animetrics [2] and ReproFace used by the FBI [3], there also exists an open source library which allows one to generate a 3D face model from arbitrary 2D facial images. This library, called Candide mikael,AHL01, may allow academia and industry to significantly improve the matching performance of their algorithms in video surveillance applications.

Candide also makes it possible to generate facial images from partially visible faces. A possible situation is the case where the surveillance team has a video sequence with the subject's face partially covered. The single video frame can then be used to generate new poses, by adjusting the face mask to the face's angle appearing in the video, and Candide can generate new poses from this information. some mark, like a scar, Candide will not generate this mark on the unseen face part. These new poses can be added to the Machine Learning (ML) algorithm to improve its accuracy to detect that particular subject.

Another point that deserves attention is the fact that, if only one picture per subject is available for the learning process, then it is practically impossible for a ML system to build an accurate model of the subject's face. Synthetically generating more facial images of the same subject would allow more accurate face models to be built, yielding overall improvement in face recognition performance.

In the following we summarize some background information in synthetic face generation, and describe how the Candide face mask can be useful in face recognition systems, and present a discussion on future work.

[1] CNN - "How CMU Biometrics Center Face Recognition Could Help Boston", May 1, 2013. CMU 3D Fave Modeling research: http://www.cmu-biometrics.org

[2] 2D-3D FACEngine Face Recognition Performance Based on SetPose Geometric Normalization: http://animetrics.com/face-recognition-based-on-setpose-geometric-normalization/, Making Faces ID-Ready (The world's leading 2D to 3D face biometric forensics tool, supporting 45 pose correction, making faces "ID-Ready" for any facial recognition system): https://id.ready.animetrics.com/

[3] Richard W. Vorder Bruegge, Facial Recognition and Identification Initiatives, Federal Bureau of Investigation, Biometric Consortium Conference (BCC 2010) Sept. 2010. Online: http://biometrics.org/bc2010/presentations/DOJ/vorder_bruegge-Facial-Recognition-and-Identification-Initiatives.pdf.

2 Background – Synthetic Face Generation

In synthetic face generation approaches, a 2D face image is typically mapped onto a deformable 3D face model which is then used to allow generation of synthetic 2D faces with different angles and poses. A 3D face model can be generated from one or more images automatically and then can be adjusted by manually by mapping facial features onto internal face model. Afterwards, a morphable face model can be derived by transforming the shape and texture of the 3D face model into set of vectors. Linear combination of thus obtained vector prototypes allows one to model new facial apperances and expressions.

A combination of 3D morphable models and component-based recognition has been used for building pose and illumination invariant FR systems [5]. Three input faces of each person are employed in [5] to compute morphable 3D face models, which are then used to build a large set of synthetic faces under different viewpoints and lighting conditions for training a component-based FR system. Initial database of 3D models was built with a 3D laser scanner. By morphing between the existing models in the database, pose ranges within a range of +/- 45 degrees of rotation in depth and +/- 10 degrees of rotation in the image plane can be achieved, using two illumination models for each pose.

Synthetic face cubes extracted from original face images in both frontal and 20 degrees side views are introduced in [7] based on head shapes and feature location in order to match synthetic faces. The geometric difference between the faces in a four dimensional face subspace using local Euclidean distance is used as a metric in the face space.

Recently, a morphing procedure has been proposed to create training set to design a user-specific face recognizer using combination of two parallel classifiers, one based on Gabor features and the other based on Local Binary Patterns (LBP) [8, 1]. In the morphing procedure, borderline faces are generated between each target face and random non-target faces. The morphed faces can be similar to each other, where the less morphed faces can be considered as a borderline pattern of positive training samples and the deeper morphed faces are related to borderline pattern of negative training samples.

In [9], virtual samples are constructed from a single face image using a wavelet transform. First, a 2D wavelet transform is applied to decompose a facial image into four regions in the frequency domain. Then, virtual samples are generated by rotating the image in different directions. The Principle Component Analysis (PCA) is used for classification.

Wavelet transform is a time-frequency scale transformation that is developed by the Fourier transform. Regions in the frequency domain belong to one of four regions: low-frequency region LL (approximate component) and high frequency regions, LH (horizontal component), HL (vertical component), and HH (diagonal component). For example, in a 2-level wavelet decomposition the second decomposition is computed in the LL1 region. Each face is then divided into four faces of the same dimension and the quarter size of the original image.

In [10] a method called single image subspace (SIS) is proposed for single-sample-per-person problems to represent each single image as a subspace spanned by its synthesized images. Synthesized samples are used to generate subspaces, which can be constructed in three ways: 1) from the entire extended training set, 2) from all synthesized images of the subject, or 3) from all images that passed a common filter criteria.

Several other approaches to 3D modeling and its use for face recognition in video are presented in [6].

3 Candide

One of the simple, popular, and publicly available tools to generate a 3D face model from arbitrary 2D facial images is called Candide [2, 3], created by Mikael Rydfalk at Linköping University in 1987 [2]. Candide uses a parametrized face mask that is specifically developed for model-based coding of human faces. It allows fast and computationally low generation of synthetic faces using an image of frontal face or several images with partially occluded faces captured under different poses and angles. The constructed 3D model is defined by a triangulated mesh and contains a full 3D description of vertex locations of the mesh. Candide is controlled by global and local Action Units (AUs). The global AUs correspond to the rotations around three axes. The local AUs control the mimics of the face so that different expressions can be obtained. An example of the Candide wire-frame face model is shown in Figure 1.

Having the 3D model, it is possible to use standard Computer Graphics texture-mapping techniques to synthesize as many virtual face images at novel view angles as necessary.

There are implementations of Candide available on the Linköping University website for both Windows© and Linux©, but they are outdated. The source code is implemented in C++, and the Windows© version does not compile because it has some missing files. The Linux© version compiles and generates the executable file, but it demands the user to reduce the video definition to allow the program to work. In order to have a functional program, it was necessary contact the authors of [11], and they have provided the MatLab© version, that is presented in Figure 2. The main disadvantage of this program is the fact that it is implemented in a closed architecture tool, which prevents its integration with other programs. The program in MatLab puts the Candide model onto the face image. The face alignment module adapts a 3D generic face model onto the face image to extract facial shape and texture information. From this step, it is possible to extract various positions of the face and generate different files to extract the features for learning step.

The idea to use Candide in the PROVE-IT project comes from [12, 11], which adapt Candide face generation to create new faces as input to a ML algorithm. The solution presented in [12, 11] is very similar to the problem examined in the PROVE-IT project, which is how to build a face recognition system from from one image (generally frontal picture) so that it recognize the

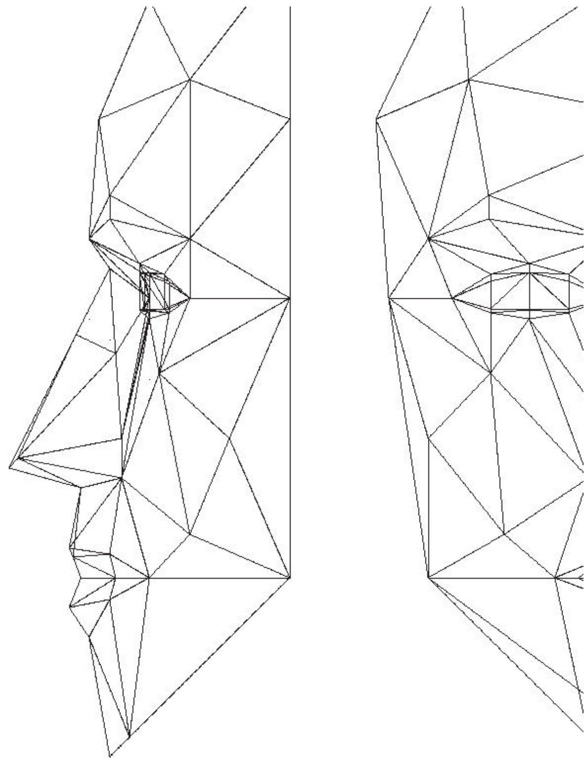

Figure 1: Candide-3 with 113 vertices and 168 surfaces.

same face under different angles of view. The approach in [12, 11] is centered on the still-to-still problem, but it gives insights on how to automate the process of creating useful data sets for still-to-video applications examined in the the PROVE-IT project.

The approach in [12, 11] used Harris detector to extract facial features and Hidden Markov Model (HMM) as a ML algorithm. The architecture of their proposed system is presented in Figure 3. Our system follows the same ideas, with the difference in the used features and ML algorithms.

4 Discussion and Future Work

The main objective of this work is to examine the applicability of open source face pose generation tools, such as Candide, for improving the performance of face recognition ML algorithms ML algorithms cannot generate reliable models from one frontal picture, which is what is normally

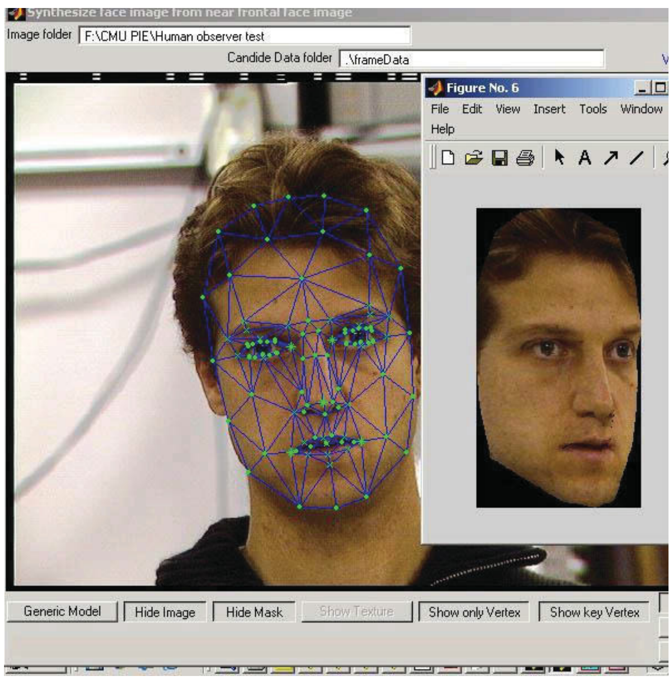

Figure 2: GUI of the synthesis module written in Matlab shows the Candide model mapped onto a face image and a synthesized face.

available in Wanted Lists. For successful face recognition, it is important from the initial frontal picture to be able to dynamically generate new facial positions, corresponding to various points-of-view, and make them available to ML algorithms, which will use them to generate better face models. This problem can be solved by using the Candide face generation tool, which builds a 3D face mesh designed to model facial pictures and which allows one to generate new poses from the original image.

The main advantage of Candide is to allow one to generate facial images from any visible position, even if the face view is partially blocked. Another capability is to allow the insertion of face expressions, like a smile, on the image. These new picture poses are useful to a ML algorithm, which can then learn different expressions and improve its recognition capabilities.

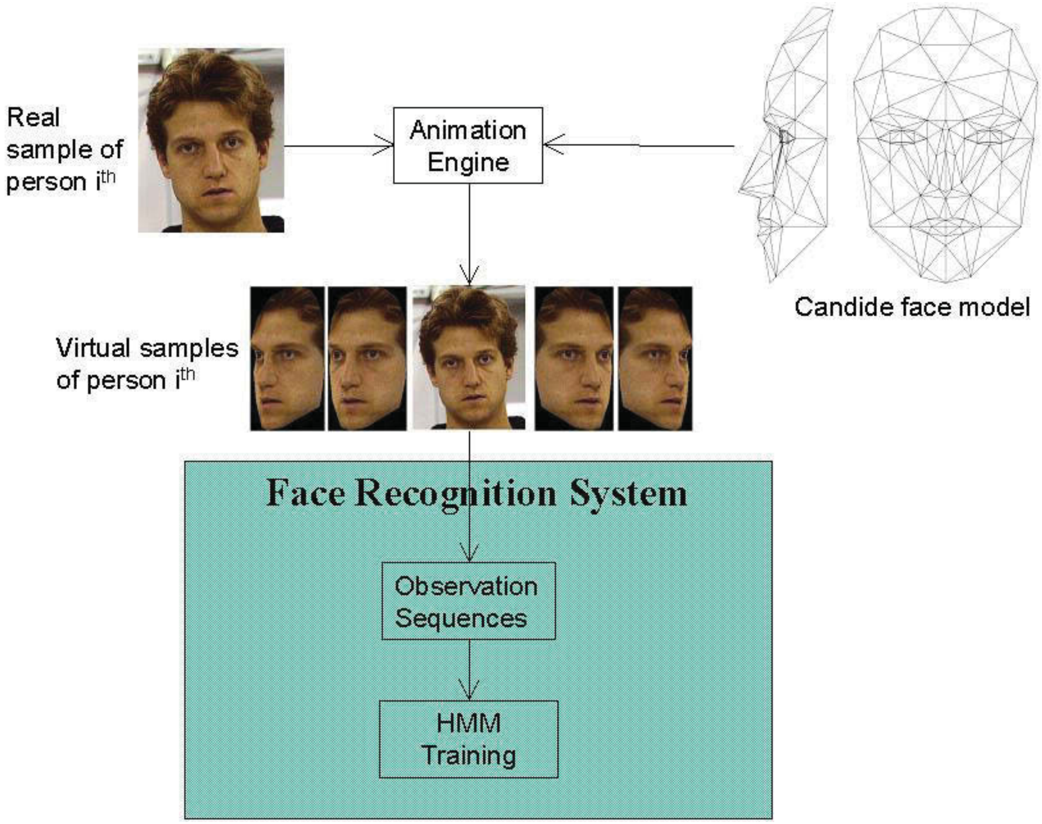

Figure 3: Illustration of HMM pose invariant system.

Candide can be used as part of a larger face recognition to allow the security personnel to use the generated face mask as part of their regular work. Their job would be to position the mask on the subject's face and then let the system generate a sequence of images corresponding to different perspectives to be used automatically by a ML algorithm.

For future work, it is necessary to translate the Candide's MatLab© implementation into a C++ library, which would allow to use the tool as to build a larger face recognition system. A standard library, like OpenCV[4], offers a natural choice, because it is an open source and free software with support to ML and Image Processing.

[4]OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code. The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution

References

[1] E. Granger, E., P. Radtke, Gorodnichy, D.: Survey of academic research and prototypes for face recognition in video. CBSA Division Report (2014).

[2] Rydfalk, M.: Candide, a parameterized face. Technical Report LiTH-ISY-I-866, Dept. of Electrical Engineering, Linköping University **1**(1), 330–338 (1987).

[3] Ahlberg, J.: Candide-3 an updated parameterised face, 2001. Online: http://www.bk.isy.liu.se/publications/LiTH-ISY-R-2326.pdf

[4] Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: ACM Press/Addison-Wesley Publishing Co., 1999, Edn., pp. 187–194.

[5] Weyrauch, B., Heisele, B., Huang, J., Blanz, V.: Component-based face recognition with 3D morphable models, in IEEE Computer Vision and Pattern Recognition Workshop on Face Processing in Video, 2004, Also in Image and Vision Computing Special Issue on Face Processing in Video Sequences (Ed. D. Gorodnichy), 2006.

[6] The First IEEE Workshop on Face Processing in Video, Track on Face modeling and matching (Ed. D. Gorodnichy), 2004, Online: http://www.visioninterface.net/fpiv04/papers.html

[7] Wilson, H.R., Loffler, G., Wilkinson, F.: Synthetic faces, face cubes, and the geometry of face space, Vision Research, 2002, 42, (27), pp. 2909-2923.

[8] Qiu, J., Zhang, Y. , Sun, J.: Face recognition in open world environment, Visual Communications and Image Processing (VCIP), 2013.

[9] Zhao, Y., Ma, Y., Ji, S.: Face recognition with single training image per person based on wavelet transform and virtual information, in IEEE Conference on Pervasive Computing Signal Processing and Applications (PCSPA), 2010.

---

image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc. OpenCV has more than 47 thousand people of user community and estimated number of downloads exceeding 7 million. The users of OpenCV include such companies as Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota and many small start-up companies around the globe. More: http://opencv.org/

[10] Liu, A., Ghosh, J., Martin, C.: A framework for analyzing skew in evaluation metrics. In: II AAAI Workshop on Evaluation Methods for Machine Mearning, pp. 22–27 (2007).

[11] Le, H.-S., Li, H.: Face identification system using single hidden markov model and single sample image per person. Proc. of the IEEE Int. Joint Conf. on Neural Networks (IJCNN) **1**(1), 330–338 (2004).

[12] Le, H.-S., Li, H.: Face identification from one single sample face image. Proc. of the IEEE Int. Conf. on Image Processing (ICIP) **1**(1), 330–338 (2004).